Pipeline overview

The pipeline takes both ChIP-seq reads (from paired-end, stranded or single-end, unstranded libraries) and a set of reference fastas as inputs, and outputs unfiltered and filtered alignment files (bam format) from the appropriate genome assembly.

When multiple fastqs are generated from a single biological replicate (in multiple sequencing runs, for example), they are concatenated before mapping.

Pipeline Restrictions

- The read length should be a minimum of 50 base pairs, though longer read lengths are encouraged.
- The sequencing platform used should be indicated.
- Replicates should match in terms of read length and run type.
- Pipeline files are mapped to either the GRCh38 or mm10 sequences: https://www.encodeproject.org/references/ENCSR425FOI/